

Spurious trend switching phenomena in financial markets

V. Filimonov^{1,a} and D. Sornette^{1,2}

¹ ETH Zurich, Department of Management, Technology and Economics, Kreuzplatz 5, 8032 Zurich, Switzerland

² Swiss Finance Institute, c/o University of Geneva, 40 blvd. du Pont d'Arve, 1211 Geneva 4, Switzerland

Received 20 December 2011 / Received in final form 2 March 2012

Published online 16 May 2012 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2012

Abstract. The observations of power laws in the time to extrema of volatility, volume and intertrade times, from milliseconds to years reported by Preis et al. (2010, 2011), are shown to result straightforwardly from the selection of biased statistical subsets of realizations in otherwise featureless processes such as random walks. The bias stems from the selection of price peaks that imposes a condition on the statistics of price change and of trade volumes that skew their distributions. For the intertrade times, the extrema and power laws results from the format of transaction data.

1 Introduction

The random walk model of Bachelier (1900) [1], later extended into the geometrical Brownian model (GBM) [2], forms a reasonable first-order approximation of the dynamics of financial market prices. The GBM constitutes the starting point for more refined modern models that take into account the stylized facts documented in the last 50 years. While the GBM is based on the generally verified absence of linear correlation of returns, many studies have shown in addition the existence of long-memory in the volatility, volatility clustering and multifractality, fat tails in the distributions of returns, correlation between volatility and volume, time-reversal asymmetry, the leverage effect, gain-loss asymmetries and many others (see for instance, [3,4]).

Recently, Preis et al. [5–11] have claimed the discovery of a new stylized fact in the form of universal power laws associated with so-called switching points. They found that local maxima of volatility and volume, and local minima of intertrade times, are reached and followed by power laws in the time to the extrema that are reminiscent of critical points in physics. The power laws are found to hold from time scales ranging from milliseconds to years.

Here, we show that these power laws are also found in the minimal random walk (they also hold as well for the GBM). They derive from the statistical method used by Preis et al. to define the switching points. Using Occam's razor, this suggests that there is no need to invoke new properties for real financial prices, since essentially all characteristics of switching points documented by Preis et al. [5–11] can be recovered in the GBM. In other words, we show that the discovery of Preis et al. is likely an artifact of their statistical analysis, which does not account for

the impact of conditioning associated with the definition of switching points on the statistical properties of financial returns. Our finding applies directly to the volatility. For volumes, one just needs to take into account the correlation between absolute price increments and volume, which is very strong at the daily time scale and weaker but yet pronounced at the tick-by-tick time scale. Concerning the dynamics of intertrade intervals, we show that it originates simply from the format of the transaction price data.

2 Analysis of volatility peaks

The local extrema considered by Preis et al. [5–11] are defined as follows. The transaction price $p(t_0)$, where t_0 is a discrete time in the interval $[0, T]$ measured in transaction number or in calendar time, is defined to be a local maximum (resp. minimum) of order Δt if there is no higher (resp. smaller) transaction price in the interval $t_0 - \Delta t \leq t \leq t_0 + \Delta t$. Independently of any assumption on the underlying generating process, this definition imposes stringent conditions on the price increments before and after the extrema. For instance, by virtue of the definition of the existence of a local maximum, the conditional expectation of the last price increment $\Delta p(t_0) = p(t_0) - p(t_0 - 1)$ leading to the maximum must be positively skewed and the conditional mean of $|\Delta p(t_0)| = |p(t_0) - p(t_0 - 1)|$ has to be larger than the unconditional mean of $|\Delta p(t)|$. This skewness also holds for earlier increments from $t_0 - 2$ to $t_0 - 1$, from $t_0 - 3$ to $t_0 - 2$, and so on, with a decreasing amplitude as one considers price increments further away from the peak. Similarly, the increment $\Delta p(t_0 + 1) = p(t_0 + 1) - p(t_0)$ is negatively skewed, with the conditional mean being smaller than the unconditional mean. This negative skewness also holds for later increments with a progressively decreasing

^a e-mail: vfilimonov@ethz.ch

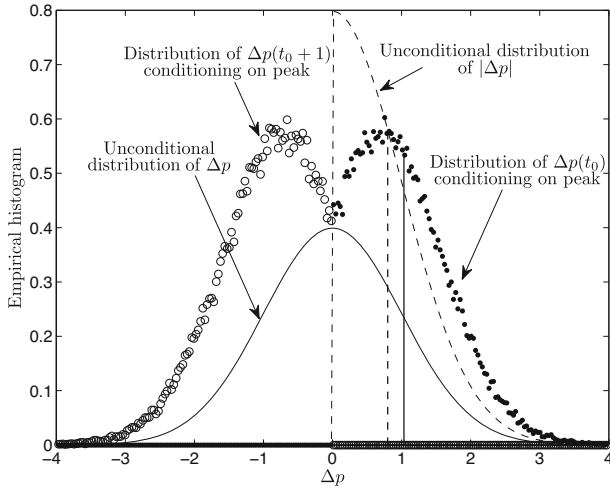


Fig. 1. Conditional and unconditional distributions of price increments (see text). The vertical lines correspond to the unconditional expectation of $|\Delta p(t)|$ (dashed line) and to the expectation of $\Delta p(t_0) = p(t_0) - p(t_0 - 1)$ conditioned on the price peak occurring at time t_0 (continuous line) for a price following the random walk model with iid Gaussian increments $N(0, 1)$. The distribution of $\Delta p(t_0) = p(t_0) - p(t_0 - 1)$ is strongly positively skewed, in particular with no negative realizations, as results from the definition of $p(t_0)$. Similarly, the distribution of $\Delta p(t_0 + 1) = p(t_0 + 1) - p(t_0)$ is strongly negatively skewed, in particular with no positive realizations, as results from the definition of $p(t_0)$.

amplitude. While these effects can be studied in details analytically for the GBM, for pedagogy, we choose to show the result of numerical simulations for the random walk model

$$p(t) = \sum_{i=0}^{t-1} \xi(i) \quad (1)$$

with Gaussian $N(0, 1)$ iid increments $\xi(i)$ (Fig. 1). The same properties as shown in Figure 1 hold for the volume and intertrade waiting times.

Preis et al. [5–11] have constructed the average trajectories of volatility, volume and intertrade waiting times as a function of time to the extrema. They have averaged twice, first, over different trend durations and, second, over different orders Δt . A positive trend duration is simply defined as the time between the last minimum to the next maximum. A negative trend duration is defined as the time between the last maximum to the next minimum. In order to be able to stack all positive and negative peaks respectively on each other, Preis et al. have normalized each trend duration to 1 by defining a dimensionless time ε taking the value 0 at the beginning of the trend and 1 at the extrema. They have found that volatility, volume and intertrade waiting times as a function of the dimensionless reduced time ε are power laws of the distance $|\varepsilon - 1|$ to the extrema.

We have performed exactly the same analysis as Preis et al. [5–11] on the random walk model (1) with Gaussian $N(0, 1)$ iid increments $\xi(i)$ and find essentially the same power law dependences as found empirically on financial

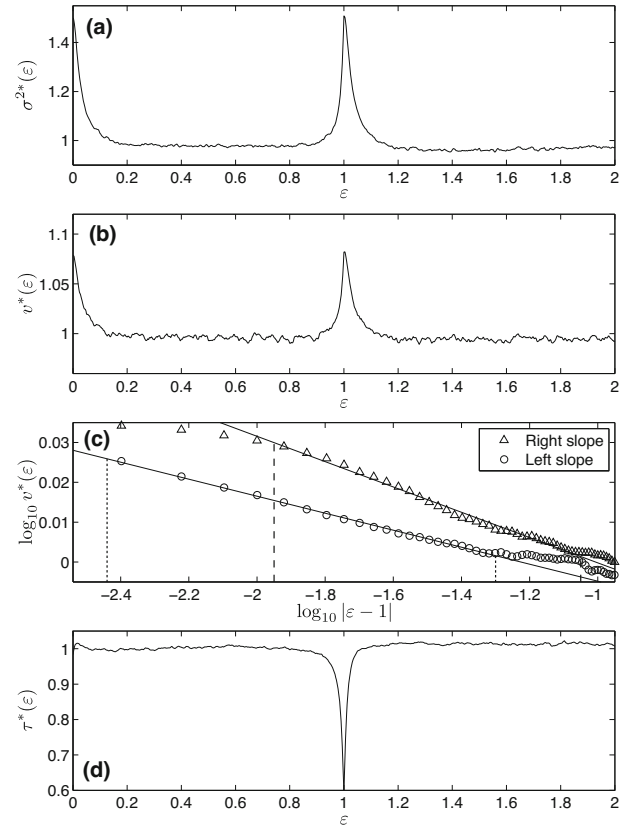


Fig. 2. (a) Financial volatility $\sigma^{2*}(\varepsilon)$ obtained from numerical simulations of the random walk model; (b) volume $v^*(\varepsilon)$ obtained from numerical simulations with the model of linear correlations between the volume and absolute price change (see text); (c) log-log plot of $v^*(\varepsilon)$ as a function of $|\varepsilon - 1|$ on both side of the peak placed at $\varepsilon = 1$. The two straight lines correspond to the power laws $v^*(\varepsilon) \sim |\varepsilon - 1|^{\beta_v}$, with $\beta_v^+ = -0.16$ in the range $10^{-1.95} < \varepsilon - 1 < 10^{-1.05}$, and $\beta_v^- = -0.12$ in the range $10^{-2.45} < 1 - \varepsilon < 10^{-1.3}$; (d) intertrade time $\tau^*(\varepsilon)$ obtained from numerical simulations with the model with non-zero probability of “walking up the book” (see text) with $p_0 = 0.5$.

time series (see Fig. 2). Figure 2a shows the average stacked peak of volatility as a function of ε for the random walk model. Note that the scale-free structure of the random walk ensures a weak dependence of this pattern on the order Δt , as also found for the real financial data. The asymmetry of the peak, as well as the two power-law decays around the peak, are found to hold also for the random walk model: a fit of the synthetic data generated with the random model with expression

$$\sigma^{2*}(\varepsilon) \sim |\varepsilon - 1|^{\beta_\sigma} \quad (2)$$

yields $\beta_\sigma^+ = -0.16$ in the range $10^{-1.95} < \varepsilon - 1 < 10^{-1.05}$, and $\beta_\sigma^- = -0.12$ in the range $10^{-2.45} < 1 - \varepsilon < 10^{-1.3}$. Note that a better model for the singular behavior for $\varepsilon \rightarrow 1$ amounts to replacing (2) by $\sigma^{2*}(\varepsilon) = a - b|\varepsilon - 1|^{\beta_\sigma}$ where a is a constant, so that the conditional volatility does not really diverge, but exhibits a finite-time singularity characterized by an infinite slope. A fit to the synthetic data usually gives $b > 0$ and $\beta_\sigma \simeq 0.7$.

The asymmetry around the peak results from (i) the selection of realizations with peaks that ensure that large positive (resp. negative) price changes are more probable before (resp. after) the peak and (ii) the asymmetry in dimensionless time ε , that is by definition fixed to the single trend for $0 \leq \varepsilon \leq 1$, but is free for $1 \leq \varepsilon \leq 2$ and may include in these values another trend, or even multiple trends, when Δt is large. This effect becomes even stronger for tick-by-tick data, where the probability of having zero price change is high and the micro trend (of given order) is likely to be followed by a plateau of constant price.

These results are robust with respect to changes in the generating price process. For instance, taking into account the existence of an atom (probability concentration) for zero price increments, or of a negative auto-correlations of price increments at lag one due to bid-ask bounce and other effects do not change qualitatively the results shown in Figure 2. Moreover, the presence of the peak in volatility is robust to detrending of the initial time series (by removing the linear/polynomial trend or using moving averages). As described above, the peak stems from the bias in statistics due to the conditioning. As long as the time series is not degenerated (in particular, the support of the probability distribution of increments $\xi(i)$ should contain more than one point in each half-axis: $\xi > 0$ and $\xi < 0$), conditioning on peaks will bias the average towards bigger increments, and thus skew the distribution of $\Delta p(t_0)$, resulting in the peak of volatility. However, the quantitative values of the parameters, such as the exponents β_σ^+ and β_σ^- and the ranges of the power-law scaling, strongly depend on the particular model and on the methods of preprocessing of the data (e.g. filtering or detrending).

3 Analysis of volume and intertrade interval peaks

Since the random walk model does not include any volume or intertrade intervals, we need to enrich it slightly to match the observations of Preis et al. [5–11]. We start with the well-established empirical fact that volume is correlated with the absolute price change at many scales. Indeed, as Preis et al. have noted [5–11], there are almost no correlation between signed price increments $\Delta p(t)$ and volume $v(t)$, due to the almost symmetric distribution of the $\Delta p(t)$. But, at the same time, it is well known that the correlation between absolute price increments $|\Delta p(t)|$ and volume is very strong at the daily scale (see for instance the early review [12]) and weaker but clearly pronounced at the transactions scale. Here, it should be noted that studies of transaction data are not common because the transaction price is subjected to bid-ask bounce and thus does not reflect well market price moves.

For illustration, let us take the example of transaction data of December 2011 futures on the DAX index over the period 07/11/2011–06/12/2011 (1 892 243 transactions in total) obtained from Bloomberg Historical Intraday Tick database. For this dataset, we find that the Pearson correlation coefficient between signed price increments $\Delta p(t)$

and volume $v(t)$ is equal to 1.02% with the 95% confidence interval [0.88%, 1.16%] clearly excluding the null hypothesis of a zero value. The correlation coefficient between absolute price increments $|\Delta p(t)|$ and volume $v(t)$ is more than 10 times stronger and equal to 15.7% with its 95% confidence interval equal to [15.6%, 15.9%].

Due to the correlation between volume and absolute price change, the structure of the volume dynamics reproduces qualitatively that of the absolute price increments. Assuming again that the price follows a random walk (1), we account for the correlation between volume and absolute price change by specifying the following process for the volume:

$$v(t) = \left| |\Delta p(t)| + \sigma_\mu \mu(t) \right|, \quad (3)$$

where $\mu(t)$ is an iid Gaussian noise $N(0, 1)$ independent of $p(t)$ and σ_μ is a coefficient that controls the amplitude of the correlation between $|\Delta p(t)|$ and $v(t)$. Figure 2b shows the average stacked volume dynamics $v^*(\varepsilon)$, conditional on price peak, as a function of ε for the random walk model with (3) and correlation coefficient of 0.2. One can clearly observe the same asymmetric peak as for the volatility shown in Figure 2a. Around the peak of volume, which coincides with that of price, Figure 2c shows a log-log plot of the volume $v^*(\varepsilon)$ as a function of $|\varepsilon - 1|$ on both side of the peak placed at $\varepsilon = 1$. This log-log plot demonstrates the existence of two power laws, according to $v^*(\varepsilon) \sim |\varepsilon - 1|^{\beta_v}$, with $\beta_v^+ = -0.16$ in the range $10^{-1.95} < \varepsilon - 1 < 10^{-1.05}$, and $\beta_v^- = -0.12$ in the range $10^{-2.45} < 1 - \varepsilon < 10^{-1.3}$. As for the conditional volatility around the peak, we note that a better model is $v^*(\varepsilon) = a' - b'|\varepsilon - 1|^{\beta_v}$ with $b' > 0$ and $\beta_v < 1$.

As shown in Figure 2d, we are also able to reproduce the negative peak of the average intertrade intervals τ as a function of ε , conditioned on the existence of a price peak, and its approximate power law dependence in terms of $|\varepsilon - 1|$ on both side of the peak. For this, we consider the same random walk model (1) with discrete price increments $\xi(i)$ and we model intertrade intervals $\tau(t)$ as a mixture of (i) iid exponentially distributed random variables reflecting a Poisson process for the order flow and (ii) an atom at $\tau = 0$ with probability mass p_0 , which accounts for those times t_c 's such that the price increment $\Delta p(t_c)$ has the same sign as the previous nonzero increment $\Delta p(t'_c)$ (this means that the price moved in the same direction as before). This is a simple toy model of the way that the transaction data is organized. In the order-driven exchanges such as the Chicago Mercantile Exchange and the Frankfurt Stock Exchange, which were analyzed by Preis et al. [5–11], the matching of orders is performed according to a real-time “walk” within the order book. When one submits a market order to buy or sell, it is immediately executed at the best possible price. But when the submitted order is so large that it could not be executed at one price (due to insufficient supply), a part of it is executed at the best next price and the rest – at the second best next price, then at the third next best price – and so on until the whole order is executed. This effect is known as “walking up the book”. These executions correspond to

different transactions, since they are performed at different prices but, since they are triggered by the same order, all of them have identical time stamp in the transaction log-file. At the same time, such sequences of transactions are moving price significantly, especially if more than two levels in the order book are involved. Therefore, the logfile of transaction records a finite fraction of trades with non-zero price changes and zero intertrade time interval. As demonstrated by Figure 2d, this effect is sufficient to explain the negative peak in the time series of the intertrade times, reported initially by Preis et al. [5–11].

4 Extension of the analysis beyond the random walk model

Finally we would like to notice that the range of the apparent power laws and their exponents obtained in the random walk model (1) do not match quantitatively the results reported by Preis et al. [5–11]. However, supplementing the random walk model with another stylized fact of real financial time series, it is possible to reproduce precisely both the quantitative values of the exponents and the range of scales over which the power laws hold. To illustrate this result, we will use the quasi-multifractal process [13,14] that accounts for the heavy tails of price increments and the long memory of absolute returns. This model represents price $p(t)$ as

$$p(t) = \sum_{i=0}^{t-1} \xi(i) e^{\omega(i)}, \quad (4)$$

where $\xi(i)$ are i.i.d. Gaussian $N(0,1)$ variables and $\omega(i)$ are independent of $\xi(i)$ Gaussian random variables with zero mean and covariance matrix

$$\mathbb{B}_{i,j} = \frac{\sigma^2 \varphi}{2} \int_0^\infty \frac{dx}{((1+x)(1+x+|i-j|))^{\varphi+1/2}}. \quad (5)$$

Figure 3 presents an example of the simulations of the quasi-multifractal model (4)–(5) for values of the parameters ($\varphi = 0.1$ and $\sigma^2 = 5$) that replicate the power laws found in the volatility patterns of the S&P 500 Stocks documented in references [5,8,11], with exponents $\beta_\sigma = -0.46$ in the range $10^{-1.7} < \varepsilon - 1 < 10^{-0.4}$.

5 Conclusion

In summary, we have shown that the definition of price peaks imposes a condition on the statistics of price change and of trade volumes that skew their distributions sufficiently to explain the occurrence of power laws in the time to these peaks, even in the simplest possible model, the random walk. Though the minimal random walk model could not reproduce precisely the exact exponents and ranges of power laws reported by Preis et al. [5–11], more elaborated models allow to match these values quantitatively. This statement was illustrated with the quasi-multifractal model [13,14] that accounts for the long-term

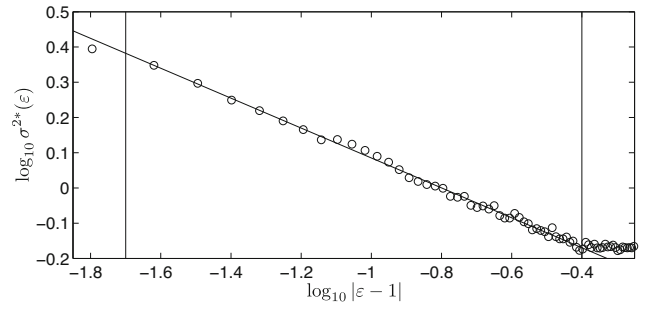


Fig. 3. Log-log plot of the volatility $\sigma^{2*}(\varepsilon)$ obtained from numerical simulations of the quasi-multifractal model (4)–(5) introduced in references [13,14] with parameters $\varphi = 0.1$ and $\sigma^2 = 5$. The straight line corresponds to the power law with the exponent $\beta_\sigma = -0.45$.

memory and heavy tailed statistics of real price returns. For the intertrade times, we have shown that the extrema and power laws can result from the format of transaction data. We are thus led to conclude that there is no new “switching” phenomenon, as the power laws are straightforward consequences of the selection of biased statistical subsets of realizations in otherwise featureless processes. In the switching phenomena reported by Preis et al. [5–11], there is no more than statistical conditioning and some correlations.

Conditional statistics are in general far more complicated than unconditional statistics and may lead to quite non-intuitive results, as we have illustrated here. Nowadays, misinterpretations of results obtained from conditional statistics seem to become more and more frequent in research (especially in complex system sciences) that relies heavily on data mining and data analysis (see for instance spurious claims of novel phenomena in the statistical physics of earthquakes that have been debunked in [15–17]). The easy access to very large data sets and the use of large computational power allow researchers to pose sophisticated questions to data by developing extremely complex statistical metrics based on conditional probabilities and averages. Being strongly dependent on the posed conditions, such metrics may lead to apparent nontrivial results, even in structureless systems, such as random walks or GBM. The literature has several remarkable examples on par with the spurious switching phenomena that we have analyzed in the present paper. Let us for instance mention another example of misinterpretation of the conditional statistics also belonging to econophysics: Plerou et al. [18] interpreted conditional probability distributions of price changes as the evidence for phase transitions in financial markets, when the same patterns could be found in simple counter-examples using the trivial GBM [19]. Paying a lot of attention to the careful implementation of the novel proposed statistical metric, one could easily neglect the influence of the complexity of the question asked on the results, thus increasing the possibility of mistaking spurious effects for reality. The impact of the complexity and distortions introduced by the conditioning method of analysis should be always kept in mind. A general approach should follow

the treatment presented here and in [19] (see also the admonitions in [15–17]), namely always test the effect of conditioning and of new metrics on known data, for instance standard processes whose properties are well-known. Only in this way can one disentangle what in the results derives from the conditioning analysis and what is a real structure of the data.

We are grateful to G. Harras for useful discussions while preparing this manuscript.

References

1. L. Bachelier, Ann. Sci. Éc. Norm. Supér. 3^e Sér. **17**, 21 (1900)
2. P. Cootner, *The Random Character of Stock Market Prices* (MIT Press, Cambridge, 1965)
3. J.-P. Bouchaud, M. Potters, *Theory of financial risks: from statistical physics to risk management*, 1st edn. (Cambridge University Press, Cambridge, 2000), p. 232
4. R. Cont, Quant. Financ. **1**, 223 (2001)
5. T. Preis, Eur. Phys. J. Spec. Top. **194**, 5 (2011)
6. T. Preis, J.J. Schneider, H.E. Stanley, PNAS **108**, 7674 (2011)
7. T. Preis, H.E. Stanley, Physics World **24**, 29 (2011)
8. T. Preis, H.E. Stanley, J. Stat. Phys. **138**, 431 (2010)
9. H.E. Stanley, S.V. Buldyrev, G. Franzese, S. Havlin, F. Mallamace, P. Kumar, V. Plerou, T. Preis, Physica A **389**, 2880 (2010)
10. T. Preis, H.E. Stanley, Trend Switching Processes, in Financial Markets, in *Econophysics Approaches to Large-Scale Business Data and Financial Crisis*, edited by M. Takayasu, T. Watanabe, H. Takayasu, (Springer, Tokyo, Japan, 2010), pp. 3–26
11. T. Preis, H.E. Stanley, Bulletin of the Asia Pacific Center for Theoretical Physics **23**, 18 (2009)
12. J. Karpoff, J. Financ. Quant. Anal. **22**, 109 (1987)
13. A. Saichev, D. Sornette, Phys. Rev. E **74**, 011111 (2006)
14. A. Saichev, V. Filimonov, JETP Lett. **87**, 506 (2008)
15. A. Helmstetter, D. Sornette, Phys. Rev. Lett. **92**, 129801 (2004)
16. M. Werner, D. Sornette, Phys. Rev. Lett. **99**, 179801 (2007)
17. A. Saichev, D. Sornette, Phys. Rev. Lett. **97**, 078501 (2006)
18. V. Plerou, P. Gopikrishnan, H.E. Stanley, Nature **421**, 130 (2003)
19. M. Potters, J.-P. Bouchaud, Comment on: Two-Phase Behaviour of Financial Markets (2003), [arXiv:cond-mat/0304514v1](https://arxiv.org/abs/cond-mat/0304514v1)